Convergence results in multitype, multivariate, branching random walk

J. D. Biggins,
The University of Sheffield, UK

Department of Probability and Statistics,
The University of Sheffield,
Sheffield, S3 7RH,
U.K.
j.biggins@sheffield.ac.uk

A. Rahimzadeh Sani,
Teacher Training University of Tehran, Iran

Department of Mathematics,
Teacher Training University of Tehran,
49 Mofatteh Avenue, Tehran, 15614,
Iran.
rahimsan@saba.tmu.ac.ir

Keywords and phrases: large deviations; local limit theorem; uniform convergence; multitype branching random walk; Laplace transforms; martingales; Perron-Frobenius; non-negative matrices
MS2000 subject classification. Primary 60J80 Secondary 15A48; 60F10

Abstract

A multitype branching random walk on d-dimensional Euclidian space is considered. The uniform convergence, as n goes to infinity, of a scaled version of the Laplace transform of the point process given by the nth generation particles of each type is obtained. Similar results in the one-type case, where the transform gives a martingale, have been obtained in Biggins (1992) and Barral (2001). This uniform convergence of transforms is then used to obtain limit results for numbers in the underlying point processes. Supporting results, which are of interest in their own right, are obtained on the normal approximation of multivariate distributions and on 'Perron-Frobenius theory' for matrices that are smooth functions of a variable $\lambda \in L$ and non-negative when $\lambda \in L_- \subset L$.

1 Introduction

We consider the multitype branching random walk on the d-dimensional Euclidian space, $\mathbb{R}^d$. The process starts with a single particle located at the origin. This particle produces daughter particles, which are scattered in $\mathbb{R}^d$, to give the first generation. These first generation particles produce daughter particles to give the second generation, and so on. As usual in branching processes, the nth generation particles reproduce independently of each other. Particles in this process have one of p types. For each type i there is a vector of point processes $(Z_{i1}, Z_{i2}, \ldots, Z_{ip})$. Then, when a type i particle reproduces, the positions of its daughter particles of the various types, relative to the parent's position, are given by a copy of $(Z_{i1}, Z_{i2}, \ldots, Z_{ip})$. The one-type branching random walk has an extensive literature. The multitype extension has received less attention, but discussion of it can be found in Mode (1971), Biggins (1976, 1996), Bramson *et al.* (1992) and Kyprianou and Rahimzadeh Sani (2001).

We keep i, j and k, for the types, drawn from $\{1, \ldots, p\}$. Let μ_{ij} be the intensity measure of Z_{ij}. We assume throughout that

$$\text{there is a } \vartheta \in \mathbb{R}^d \text{ such that } \max_{i,j} \int_{\mathbb{R}^d} e^{-\vartheta' x} \mu_{ij}(dx) < \infty, \tag{1.1}$$

where $\vartheta' x = \sum_i \vartheta_i x_i$ is the usual inner product of vectors. This condition is enough to ensure that convolutions of the μ_{ij} produce well-defined measures.

Let Z_{ij}^n be the point process giving the positions of the type j particles in generation n when the initial ancestor is of type i. Then Z_{ij}^1 is distributed like Z_{ij}. The first objective here is to obtain the asymptotic behaviour, as $n \to \infty$, of the Laplace transform of Z_{ij}^n. Before the theorems can be stated, further notation about intensity measures and their transforms is needed.

Define $\mu_{ij}^{n\star}$ inductively by

$$\mu_{ij}^{(n+1)\star} = \sum_{k=1}^{p} \mu_{ik} * \mu_{kj}^{n\star}$$

where "$*$" is ordinary convolution of measures. It is easy to confirm, by induction on n, that the point process Z_{ij}^n has the intensity measure $\mu_{ij}^{n\star}$. Furthermore the counts of the numbers of each type in each generation, given by $(Z_{i1}^n(\mathbb{R}^d), \ldots, Z_{ip}^n(\mathbb{R}^d))$, is a multitype Galton-Watson process, which is discussed in Athreya and Ney (1972), for example. A matrix A of non-negative entries is called positive regular when for some positive integer n, A^n has all its entries strictly positive. The multitype Galton-Watson process is positive regular when the matrix $(P(Z_{ij}(\mathbb{R}^d) > 0))$ is positive regular. Throughout, the embedded Galton-Watson process is assumed to be positive regular. This Galton-Watson process is supercritical when the largest eigenvalue of its mean matrix, $(EZ_{ij}(\mathbb{R}^d))$, exceeds one. (For a positively regular process this eigenvalue will be infinite when the mean matrix has any infinite entries.) A supercritical process survives with positive probability. One of the conditions of our theorems will imply that the process is supercritical.

The d-dimensional complex space $\mathbb{C}^d$ is equipped with the maximum metric. Hence, for $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d) \in \mathbb{C}^d$, we have $d(x, y) = \max\{|x_i - y_i| : i = 1, \ldots, d\}$ where $|x_i - y_i|$ is the usual absolute value in $\mathbb{C}$. Let $B(x, r)$ be the open ball centred at x of radius r using this metric and let $\overline{B}(x, r)$ be its closure. (Later we will also need

$S(x,r)$ and $\overline{S}(x,r)$ for the open and closed balls in $\mathbb{R}^d$.) In all that follows, we keep the letters θ and η for the real and imaginary parts of $\lambda \in \mathbb{C}^d$, so that, with this convention, $\lambda = \theta + \mathbf{i}\eta$. For any $A \subset \mathbb{C}^d$ let A_- be its intersection with $\mathbb{R}^d$; thus

$$A_- = \{\lambda \in A : \lambda = \theta + \mathbf{i}\eta \text{ with } \eta = 0\}.$$

Define the Laplace transforms $m_{ij}(\lambda)$, for $\lambda \in \mathbb{C}^d$, by

$$m_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda' x} \mu_{ij}(dx) \ \left(= E \int_{\mathbb{R}^d} e^{-\lambda' x} Z_{ij}(dx)\right).$$

Then $|m_{ij}(\lambda)| \le m_{ij}(\theta)$ and, by Hölder's inequality, $\{\theta \in \mathbb{R}^d : m_{ij}(\theta) < \infty\}$ is a convex set. Let int$A$ be the interior of A and let

$$L = \bigcap_{i,j} \text{int}\{\lambda = \theta + \mathbf{i}\eta \in \mathbb{C}^d \ : \ m_{ij}(\theta) < \infty\}.$$

Assumption (1.1) is now strengthened to:

$$L \text{ is non-empty.}$$

Then L is a non-empty, open, convex subset of $\mathbb{C}^d$ and each $m_{ij}(\lambda)$ is analytic in $\lambda \in L$. Let $M(\lambda)$ be the matrix given by $M(\lambda) = (m_{ij}(\lambda))$ and let $M^n(\lambda)$ be its nth power, with (i,j)th entry $m^n_{ij}(\lambda)$. Furthermore, let

$$\mathcal{M}^n_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda' x} Z^n_{ij}(dx).$$

Then

$$m^n_{ij}(\lambda) = \int_{\mathbb{R}^d} e^{-\lambda' x} \mu^{n\star}_{ij}(dx) = E\mathcal{M}^n_{ij}(\lambda).$$

By analogy with the case when the entries are non-negative, the eigenvalue ρ of M is called the *maximum-modulus eigenvalue* if it is a simple root of $\det[zI - M]$ and its modulus is strictly larger than that of all other roots. Note that when a maximum-modulus eigenvalue exists it is automatically unique.

When $\vartheta \in L_-(= L \cap \mathbb{R}^d)$, the entries of $M(\vartheta)$ are finite non-negative real numbers. The positive regularity of the embedded Galton-Watson process easily implies that $M(\vartheta)$ is then positively regular. This means that the following extension of the Perron-Frobenius, which is discussed in the final section, applies to M. In essence, in the present context, it says that Perron-Frobenius properties extend smoothly to complex arguments that are near to real ones. Part (iii) gives an asymptotic estimate, equation (1.2), of $M^n(\lambda)$, as $n \to \infty$. The key point about (1.2) is that it gives a uniform bound on the rate of convergence in a suitable neighbourhood. This uniformity has no direct parallel in the Perron-Frobenius Theorem and needed a little care to establish.

Theorem 1 *Suppose the $p \times p$ matrix $M = (m_{ij})$ of functions defined on the open set $L \subset \mathbb{C}^d$ satisfies the following condition: for all i and j, $m_{ij}(\lambda)$ are analytic functions in $\lambda \in L$, and for all $\tilde{\lambda} \in \tilde{L} \subset L$, $M(\tilde{\lambda})$ is positive regular.*

Then there is an open set $\Omega \subset L$ containing $\tilde{L}$ such that for $\lambda \in \Omega$ the following hold.

(i) $M(\lambda)$ has a unique maximum-modulus eigenvalue, $\rho(\lambda)$, that is analytic in λ.

(ii) The left and right eigenvectors associated with $\rho(\lambda)$, $u(\lambda)$ and $v(\lambda)$, normalised so that $\sum_{i=1}^{p} u_i(\lambda) = 1$ and $\sum_{i=1}^{p} u_i(\lambda)v_i(\lambda) = 1$, are analytic in λ and, for all i, $u_i(\lambda) \neq 0$ and $v_i(\lambda) \neq 0$.

(iii) For any $\vartheta \in \tilde{L}$ there is a neighbourhood, B, within Ω and containing $B(\vartheta, \delta)$ for some $\delta > 0$, and constants $K < \infty$ and $\gamma \in (0,1)$ such that, for all n, i and j,

$$\sup_{\lambda \in B} \left| \rho(\lambda)^{-n}(M(\lambda)^n)_{ij} - v_i(\lambda)u_j(\lambda) \right| \leq K\gamma^n. \tag{1.2}$$

All the notation introduced in Theorem 1 is now applied to M, the matrix of Laplace transforms, with $\tilde{L} = L_-$. Thus M has maximum-modulus eigenvalue ρ and eigenvectors u and v defined on the set Ω, with the properties described in Theorem 1. For all $\lambda \in \Omega$, $n = 1, \ldots$, and i, j, define

$$\mathcal{W}_{ij}^n(\lambda) = \frac{v_j(\lambda)}{v_i(\lambda)} \rho(\lambda)^{-n} \mathcal{M}_{ij}^n(\lambda) = \frac{v_j(\lambda)}{v_i(\lambda)} \rho(\lambda)^{-n} \int_{\mathbb{R}^d} e^{-\lambda' x} Z_{ij}^n(dx), \tag{1.3}$$

which is a scaled version of the Laplace transform of the point process Z_{ij}^n. Establishing the uniform convergence of such transforms in the one-type case is an integral part of the approach to results on the distribution of the points described in Biggins (1992). For the multitype case, results on the distribution of points have been obtained in Bramson *et al.* (1992) when particles are confined to the integer lattice of $\mathbb{R}$. More precise results, of the kind given in Biggins (1992), should hold and should not be limited to the lattice case. One such result is given here in Theorem 7 below.

Let $\mathcal{F}_n$ be the σ-algebra that contains all information on the multitype branching random walk up to generation n. In the one-type case, (1.3) defines a martingale with respect to $\mathcal{F}_n$, which makes aspects of the study simpler. Here, it is

$$\mathcal{W}_i^n(\lambda) = \sum_j \mathcal{W}_{ij}^n(\lambda) = \sum_j \frac{v_j(\lambda)}{v_i(\lambda)} \rho(\lambda)^{-n} \int_{\mathbb{R}^d} e^{-\lambda' x} Z_{ij}^n(dx) \tag{1.4}$$

that is a martingale with respect to $\mathcal{F}_n$ (see Lemma 2.1). Our approach to the convergence of $\{\mathcal{W}_{ij}^n(\lambda)\}$ involves condsidering the convergence of the martingale $\{\mathcal{W}_i^n(\lambda)\}$ first. Note that, given any sample path of the process, each $\mathcal{W}_{ij}^n(\lambda)$ and $\mathcal{W}_i^n(\lambda)$ is analytic in $\lambda \in \Omega$.

We need to introduce certain sets that will be used to define when convergence occurs. For $\alpha \in (1, 2]$, let

$$\Omega_\alpha^2 = \left\{ \lambda \in \Omega : \alpha\lambda \in \Omega,\ \frac{\rho(\alpha\theta)}{|\rho(\lambda)|^\alpha} < 1 \right\}, \tag{1.5}$$

$$\Omega_\alpha^3 = \operatorname{int} \left\{ \lambda = \theta + \mathbf{i}\eta \in \Omega : \max_i \{E[\mathcal{W}_i^1(\theta)^\alpha]\} < \infty \right\}, \tag{1.6}$$

$$\Lambda_\alpha = \Omega_\alpha^2 \cap \Omega_\alpha^3 \quad \text{and} \quad \Lambda = \bigcup_{1<\alpha\leq 2} \Lambda_\alpha.$$

These are all open sets in $\mathbb{C}^d$. Error estimates we derive involve $(\rho(\alpha\theta)/|\rho(\lambda)|^\alpha)^n$. Hence $\lambda \in \Omega_\alpha^2$ ensures that such bounds decay quickly with n. Obviously, $\lambda \in \Omega_\alpha^3$ imposes a moment condition. The approach requires λ to satisfy both these conditions for the same $\alpha \in (1, 2]$, which leads to the definition of Λ_α and then Λ.

The process is supercritical, that is, $\rho(0) > 1$, whenever there is a $\theta \in \Omega_\alpha^2$ for some $\alpha \in (1, 2]$. To check this, note first that $\log \rho(\theta)$ is a convex function, see Kingman (1961), Miller (1961) or Seneta (1973, Theorem 3.7). Hence

$$\frac{(\alpha - 1)}{\alpha} \log \rho(0) + \frac{1}{\alpha} \log \rho(\alpha\theta) \geq \log \rho(\theta),$$

and $\log(\rho(\alpha\theta)/\rho(\theta)^\alpha) < 0$ when $\theta \in \Omega_\alpha^2$. Then $(\alpha - 1) \log \rho(0) > 0$ and so $\rho(0) > 1$.

Theorem 2 *Let $\alpha \in (1, 2]$ and $\lambda = \theta + \mathbf{i}\eta \in \Lambda_\alpha$. Then $\{\mathcal{W}_i^n(\lambda)\}$ converges almost surely and in αth mean, as $n \to \infty$, for each i.*

This gives the following result on mean convergence as an immediate conseqeunce.

Corollary 1 *Suppose $\lambda \in \Lambda$. Then $\{\mathcal{W}_i^n(\lambda)\}$ converges almost surely and in mean, as $n \to \infty$, for each i.*

The mean convergence of $\{\mathcal{W}_i^n(\lambda)\}$ for real λ is also discussed in Bramson *et al.* (1992), Kyprianou and Rahimzadeh Sani (2001) and, rather briefly, in Biggins and Kyprianou (2004).

Theorem 3 *Assume $\Lambda_- \neq \emptyset$. Then there is an open set Γ, with $\Lambda_- \subset \Gamma \subset \Lambda$, such that the martingale $\{\mathcal{W}_i^n(\lambda)\}$ converges uniformly in any compact subset of Γ, almost surely and in mean, as $n \to \infty$, for each i. Then the limit $\mathcal{W}_i(\lambda)$ is analytic in $\lambda \in \Gamma$.*

Building on these convergence results for the martingale, analogous results will be obtained for $\{\mathcal{W}_{ij}^n(\lambda)\}$.

Theorem 4 *Let $\alpha \in (1, 2]$ and assume that $(\Lambda_\alpha)_- \neq \emptyset$. Then there is an open set Γ_1 with $(\Lambda_\alpha)_- \subset \Gamma_1 \subset \Lambda_\alpha$, such that for all $\lambda \in \Gamma_1$, $\{\mathcal{W}_{ij}^n(\lambda)\}$ converges to $u_j(\lambda)v_j(\lambda)\mathcal{W}_i(\lambda)$, almost surely and in αth mean, as $n \to \infty$, where $\mathcal{W}_i(\lambda)$ is the limit of the martingale $\{\mathcal{W}_i^n(\lambda)\}$ as $n \to \infty$.*

Theorem 5 *Assume $\Lambda_- \neq \emptyset$. Then there is an open set Γ_2, satisfying $\Lambda_- \subset \Gamma_2 \subset \Lambda$, such that for all i, j, $\{\mathcal{W}_{ij}^n(\lambda)\}$ converges uniformly in any compact subset of Γ_2, almost surely and in mean, as $n \to \infty$, to the random variable $u_j(\lambda)v_j(\lambda)\mathcal{W}_i(\lambda)$, where $\mathcal{W}_i(\lambda)$ is the limit of the martingale $\{\mathcal{W}_i^n(\lambda)\}$ as $n \to \infty$.*

The martingales $\mathcal{W}_i^n(\lambda)$ and their components $\mathcal{W}_{ij}^n(\lambda)$ are only defined on the set Ω introduced in Theorem 1. However, to move from information on transforms to information on the associated measures a result on the behaviour of $\mathcal{M}_{ij}^n(\lambda)$ is needed for $\theta \in L_-$ but where λ need not be in Ω.

The branching random walk is *strongly non-lattice* when it is positively regular and, for some (k, l) and for some $\theta \in \{\vartheta : m_{kl}(\vartheta) < \infty\}$

$$\left| \frac{m_{kl}(\theta + \mathbf{i}\eta)}{m_{kl}(\theta)} \right| = 1 \quad \text{only when} \quad \eta = 0. \tag{1.7}$$

This follows the usage of 'strongly non-lattice' in (1.64) in Bhattacharya (1977) rather than that in Stone (1965).

Theorem 6 *For a strongly non-lattice branching random walk, for any set $\mathcal{K} \subset \Lambda_-$ that is compact in $\mathbb{R}^d$ and any $\epsilon \in (0,1)$ there is an $\varepsilon < 1$ such that*

$$\varepsilon^{-n} \sup_{i,j} \sup_{\theta \in \mathcal{K}} \sup_{\epsilon \le |\eta| \le \epsilon^{-1}} \left| \frac{\mathcal{M}^n_{ij}(\theta + \mathbf{i}\eta)}{\rho^n(\theta)} \right| \to 0 \quad \text{as } n \to \infty \tag{1.8}$$

almost surely.

Between them, Theorems 5 and 6 provide strong enough information on the behaviour of the transforms $\mathcal{M}^n_{ij}(\lambda)$ to develop good estimates of the associated measures. In the one type case this was done in Biggins (1992) by relying heavily on the corresponding results for sums of independent identically distributed random variables, obtained in Stone (1967). Suitable results for Markov additive processes, which are the analogue of independent identically distributed random variables here, are not available and so the treatment has to go back to the general methods for such approximations. The results needed, which give an approximation to a measure based only on a few of its characteristics, are contained in Theorems 9 and 10 in Section 5. That section is independent of the rest of the paper and of more general interest. We illustrate there how these two theorems connect with local large deviation results in Chaganty and Sethuraman (1993) and Stone (1967).

It is worth noting that in this study the measures of interest, Z^n_{ij}, are atomic. Hence if (1.7) were replaced by the stronger, but easier to work with, Cramér's condition, (1.36) in Bhattacharya (1977), that condition could not transfer to the Z^n_{ij}.

The assumption that $\log \rho$ is strictly convex will be imposed, which amounts to saying that the branching random walk is truly d-dimensional, but we do not demonstrate this. For $\theta \in L_-$, let $(-1)^i \upsilon_i(\theta)$ be the array of ith derivatives of $\log \rho(\theta)$. Then $\upsilon_2(\theta)$ is a positive-definite matrix when $\log \rho$ is strictly convex. Let $\mathfrak{D}$ be the continuously differentiable functions and if $f \in \mathfrak{D}$ let its vector of derivatives be f'. We need functions that decay suitably at infinity and so we introduce

$$\mathfrak{G}(G) = \left\{ f \in \mathfrak{D} : \int_0^\infty \sup \{|f(x)|, |f'(x)| : |x| \ge (r-1)\} (r+1)^{d+1} dr \le G \right\}.$$

Theorem 7 *Assume $\log \rho$ is strictly convex and the process is strongly non-lattice. Let $\mathcal{K}$ be a compact subset of Λ_-. Let h be such that for all $\theta \in \mathcal{K}$ the function $e^{\theta' x} h(x)$ is in $\mathfrak{G}(G)$. Then, with $\xi(\theta) = -\theta \upsilon_1(\theta) - \log \rho(\theta)$, and C any convex set in $\mathbb{R}^d$,*

$$n^{d/2} e^{n\xi(\theta)} \int_C h(x) Z^n_{ij}(dx + n\upsilon_1(\theta)) \to \frac{v_i(\theta) u_j(\theta) \mathcal{W}_i(\theta)}{\sqrt{(2\pi)^d \det[\upsilon_2(\theta)]}} \int_C e^{\theta' x} h(x)\, dx$$

uniformly in $\theta \in \mathcal{K}$, C and G, almost surely on $\mathcal{S}$, where $\mathcal{S}$ is the survival set of the underlying Galton-Watson process.

The condition that $e^{\theta' x} h(x)$ is in $\mathfrak{G}(G)$ becomes more restrictive as $\mathcal{K}$ becomes bigger. In particular, it forces $h(x)$ to decay rapidly with $|x|$ when the origin is in the interior of $\mathcal{K}$.

Corollary 2 *Under the conditions of Theorem 7,*

$$n^{d/2} e^{n\xi(\theta)} Z^n_{ij}(C + n\upsilon_1(\theta)) \to \frac{v_i(\theta) u_j(\theta) \mathcal{W}_i(\theta)}{\sqrt{(2\pi)^d \det[\upsilon_2(\theta)]}} \int_C e^{\theta' x} dx$$

uniformly in convex $C \subset \{x : |x| \le b\}$ and $\theta \in \mathcal{K}$.

It is worth pointing out that if, for some y, $\mathcal{K}$ lies inside the half-space $\{\theta : \theta' y > 0\}$ then the result in this corollary can also hold for some C that are not bounded.

The relationships between ρ, υ_1 and ξ are the usual ones associated with large deviation and saddlepoint calculations; see Section 2.2 in Jensen (1995) for example. The assumption that ρ is strictly convex means that compact subsets of Λ_- translate, under $y = \upsilon_1(\theta)$ to compact subsets of $\text{int}\{\upsilon_1(\theta) : \theta \in L_-\}$.

Theorem 7 has the following obvious analogue for the measures $\mu_{ij}^{n\star}$, with the proof requiring only obvious changes.

Theorem 8 *Assume* $\log \rho$ *is strictly convex and the process is strongly non-lattice. Let* $\mathcal{K}$ *be a compact subset of* L_-*. Let* h *be such that for all* $\theta \in \mathcal{K}$ *the function* $e^{\theta' x} h(x)$ *is in* $\mathfrak{G}(G)$*. Then, with* $\xi(\theta) = -\theta \upsilon_1(\theta) - \log \rho(\theta)$*, and* C *any convex set in* $\mathbb{R}^d$*,*

$$n^{d/2} e^{n\xi(\theta)} \int_C h(x) \mu_{ij}^{n\star}(dx + n\upsilon_1(\theta)) \to \frac{v_i(\theta) u_j(\theta)}{\sqrt{(2\pi)^d \det[\upsilon_2(\theta)]}} \int_C e^{\theta' x} h\,(x)\, dx$$

uniformly in $\theta \in \mathcal{K}$*,* C *and* G*.*

2 Proofs of Theorems 2 and 3

For any $\lambda = \theta + \mathbf{i}\eta \in \Omega$, define the functions:

$$\bar{v}(\lambda) = \max_{i,j}\{|v_j(\lambda)/v_i(\lambda)|\}, \quad \underline{v}(\lambda) = \min_{i,j}\{|v_j(\lambda)/v_i(\lambda)|\}, \quad \nu(\lambda) = \frac{\bar{v}(\lambda)}{\underline{v}(\theta)}$$

and

$$\phi(\lambda) = \frac{\rho(\theta)}{|\rho(\lambda)|}.$$

All of these are strictly positive, continuous functions in $\lambda \in \Omega$, where Ω comes from Theorem 1. Also, since $|m_{ij}(\lambda)| \leq m_{ij}(\theta)$, we know $|\rho(\lambda)| \leq \rho(\theta)$ (Lancaster and Tismenetsky (1985, Theorem 15.2.1)) and so $\phi(\lambda) \geq 1$.

For each $\alpha \in (1, 2]$ define Ω_α by $\Omega_\alpha = \{\lambda \in \Omega : \alpha\lambda \in \Omega\}$. Then Ω_α is an open subset of Ω. For $\alpha \in (1, 2]$ and $\lambda \in \Omega_\alpha$, let

$$\nu_1(\lambda) = \frac{\bar{v}(\lambda)^\alpha}{\underline{v}(\alpha\theta)} \quad \text{and} \quad \kappa(\lambda) = \frac{\rho(\alpha\theta)}{|\rho(\lambda)|^\alpha}.$$

Then ν_1 and κ are strictly positive continuous functions in $\lambda \in \Omega_\alpha$; they depend on α, but this has been suppressed in the notation. Note that Ω_α^2, defined at (1.5), can now be written as $\Omega_\alpha^2 = \{\lambda \in \Omega_\alpha : \kappa(\lambda) < 1\}$ and is an open subset of Ω_α. Define

$$\beta(\theta) = \max_i\{E[\mathcal{W}_i^1(\theta)^\alpha]\}.$$

Then β is real valued continuous function on Ω_-, and Ω_α^3, defined at (1.6), can be written as $\Omega_\alpha^3 = \text{int}\{\lambda = \theta + \mathbf{i}\eta \in \Omega : \beta(\theta) < \infty\}$.

Let $\alpha \geq 1$. Define the αth absolute central moment of X, $\sigma^\alpha(X)$, by

$$\sigma^\alpha(X) = E\,|X - E[X]|^\alpha\,,$$

and the αth absolute central moment conditional on the σ-algebra $\mathcal{G}$ by

$$\sigma^\alpha(X \mid \mathcal{G}) = E\left[| X - E[X \mid \mathcal{G}] |^\alpha | \mathcal{G}\right].$$

Let $\{z^l_{ik;s} : s\}$ be the positions of the particles making up Z^l_{ik}. Thus $z^l_{ik;s}$ is the position of the sth particle of type k in generation l when the initial ancestor is of type i. Now, by looking at the particles in the nth generation as the $(n-l)$th generation children of the particles in generation l, we can introduce $Z^n_{kj}(\cdot|l,s)$ as the point process giving the positions of type j daughter particles in generation n descended from the sth particle with type k in generation l relative to that particle's position. Given $\mathcal{F}_l$, $\{Z^n_{kj}(\cdot|l,s) : s\}$ are independent copies of the point process Z^{n-l}_{kj}. Thus, conditional on $\mathcal{F}_l$, the random variables

$$\mathcal{W}^{n-l}_k(\lambda|l,s) = \sum_j \frac{v_j(\lambda)}{v_k(\lambda)} \frac{1}{\rho(\lambda)^{n-l}} \int_{\mathbb{R}^d} e^{-\lambda' x} Z^n_{kj}(dx|l,s) \tag{2.1}$$

are, as s varies, independent and identical copies of $\mathcal{W}^{n-l}_k(\lambda)$. Furthermore,

$$\mathcal{W}^n_i(\lambda) = \sum_k \frac{v_k(\lambda)}{v_i(\lambda)} \frac{1}{\rho(\lambda)^l} \sum_s \exp(-\lambda' z^l_{ik;s}) \mathcal{W}^{n-l}_k(\lambda|l,s). \tag{2.2}$$

Lemma 2.1 *$\{\mathcal{W}^n_i(\lambda)\}$ is a martingale with respect to $\{\mathcal{F}_n\}$.*

Proof. From Theorem 1, $E[\mathcal{W}^1_k(\lambda)] = 1$. Then decomposition (2.2) with $l = n-1$ gives the result. □

The following result, which is an extension of inequality (4) in von Bahr and Esseen (1965) to complex valued random variables, is proved in Biggins (1992); Barral (2001) provides a very pretty approach to uniform convergence in the one-type case by applying a Lemma like this to Banach-space valued random variables.

Lemma 2.2 *If $\{X_n\}$ are independent complex valued random variables with $E[X_n] = 0$, or more generally, martingale differences, then for $\alpha \in (1,2]$, there exists a constant $C > 0$, depending on α but independent of n and the sequence $\{X_n\}$, such that*

$$E\left|\sum_{j=1}^n X_j\right|^\alpha \le C \sum_{j=1}^n E\left|X_j\right|^\alpha.$$

The next lemma, which is the multitype version of Lemma 2 in Biggins (1992), gives some bounds related to the martingale $\{\mathcal{W}^n_i(\lambda)\}$. In bounding formulae like these the argument λ will often be omitted; thus, for example, ϕ is $\phi(\lambda)$.

Lemma 2.3 *Let $\alpha \in (1,2]$ and $\lambda = \theta + \mathbf{i}\eta \in \Lambda_\alpha$ and $\varphi = \nu_1 \phi^\alpha \nu^\alpha \beta$. Then there are constants c_1, c_2, and c_3, depending on α but not on λ, i or n, such that the following hold.*

(i) $\left(E\left|\mathcal{W}^{n+1}_i(\lambda) - \mathcal{W}^n_i(\lambda)\right|\right)^\alpha \le E\left|\mathcal{W}^{n+1}_i(\lambda) - \mathcal{W}^n_i(\lambda)\right|^\alpha \le c_1 \varphi \kappa^n$;

(ii) $\sigma^\alpha(\mathcal{W}^n_i(\lambda)) \le c_2 \varphi (1-\kappa)^{-1}$;

(iii) $\sum_{k=n}^\infty E\left|\mathcal{W}^{k+1}_i(\lambda) - \mathcal{W}^k_i(\lambda)\right| \le c_3 \varphi^{1/\alpha} \kappa^{n/\alpha} (1-\kappa^{1/\alpha})^{-1}$;

(iv) For $\vartheta \in (\Lambda_\alpha)_-$, there is a $\delta > 0$, such that

$$\sup\{E \mid \mathcal{W}^n_i(\lambda) \mid^\alpha : \lambda \in B(\vartheta,\delta), i, n\} < \infty.$$

Proof. The first inequality in (i) is just $E|X|^\alpha \geq [E|X|]^\alpha$. Since $E|\mathcal{W}_i^1(\lambda)|^\alpha \geq 1$, we get

$$E \mid \mathcal{W}_i^1(\lambda) - 1 \mid^\alpha \leq 4E\left|\mathcal{W}_i^1(\lambda)\right|^\alpha \leq 4\phi^\alpha\nu^\alpha\beta.$$

Using (2.2), Lemma 2.2 and that $\mathcal{W}_k^1(\lambda|n,s)$ are independent copies of $\mathcal{W}_k^1(\lambda)$ given $\mathcal{F}_n$,

$$\begin{aligned}\sigma^\alpha(\mathcal{W}_i^{n+1}(\lambda) \mid \mathcal{F}_n) &= E\left[\left|\mathcal{W}_i^{n+1}(\lambda) - \mathcal{W}_i^n(\lambda)\right|^\alpha \middle| \mathcal{F}_n\right] \\ &\leq C\sum_k \left|\frac{v_k(\lambda)}{v_i(\lambda)\rho(\lambda)^n}\right|^\alpha \sum_s \left|\exp(-\lambda' z_{ik;s}^n)\right|^\alpha E\left|\mathcal{W}_k^1(\lambda) - 1\right|^\alpha \\ &\leq c_1\nu_1\kappa^n\mathcal{W}_i^n(\alpha\theta)(\phi^\alpha\nu^\alpha\beta) = c_1\varphi\kappa^n\mathcal{W}_i^n(\alpha\theta);\end{aligned}$$

taking expectations now gives (i).

Note that, since $\lambda \in \Lambda_\alpha$, $\kappa < 1$. Now, part (ii) follows directly from summing over n in (i) and using Lemma 2.2. To prove (iii), sum the inequalities in (i).

The continuity of the functions ν, ν_1, ϕ, κ, on Λ_α and part (ii) mean that there are constants $C > 0$, depending only on α and $\delta > 0$, such that for any $\lambda \in \overline{B}(\vartheta, \delta)$, $\sigma^\alpha(\mathcal{W}_i^n(\lambda)) \leq C$. Since $E[|\mathcal{W}_i^n(\lambda)|^\alpha] \leq 2 + 2\sigma^\alpha(\mathcal{W}_i^n(\lambda))$, this implies (iv). □

Proof of Theorem 2. From part (iii) of Lemma 2.3, $E\sum_{n=0}^\infty \left|\mathcal{W}_i^{n+1}(\lambda) - \mathcal{W}_i^n(\lambda)\right|$ is finite, and so $\{\mathcal{W}_i^n(\lambda)\}$ is a Cauchy sequence almost surely. Thus as $n \to \infty$, $\{\mathcal{W}_i^n(\lambda)\}$ converges to $\mathcal{W}_i(\lambda)$ almost surely. Applying Fatou's lemma, Lemma 2.2 and Lemma 2.3(i) gives

$$\begin{aligned}E|\mathcal{W}_i(\lambda) - \mathcal{W}_i^n(\lambda)|^\alpha &\leq \liminf_{N\to\infty} E|\mathcal{W}_i^{n+N}(\lambda) - \mathcal{W}_i^n(\lambda)|^\alpha \\ &\leq C\sum_{j=0}^{k-1} E\left|\mathcal{W}_i^{n+j+1}(\lambda) - \mathcal{W}_i^{n+j}(\lambda)\right|^\alpha \\ &\leq Cc_1\varphi\frac{\kappa^n}{1-\kappa}.\end{aligned}$$

Since $\kappa < 1$, $\{\mathcal{W}_i^n(\lambda)\}$ converges in αth mean. □

The distinguished boundary (see Hörmander (1973)) of $B(x, r)$, which is a subset of the topological boundary, is defined by

$$D(x, r) = \{y \in \mathbb{C}^d : |x_s - y_s| = r; s = 1, \ldots, d\}.$$

We assume $D(x, r)$ is parameterised so that

$$D(x, r) = \{z(t) = (z_1(t), \ldots, z_d(t)) : z_s(t) = x_s + re^{\mathrm{i}t_s};\ s = 1, \ldots, d;\ t \in I\} \qquad (2.3)$$

where $I = [0, 2\pi]^d$ is a d-dimensional closed cube in $\mathbb{R}^d$. The next Lemma, which is Lemma 3 in Biggins (1992), is the key to obtaining bounds that hold uniformly.

Lemma 2.4 *If f is analytic on the open ball $B = B(x, 2r')$ with $r < r'$ then*

$$\sup\{|f(\lambda)| : \lambda \in B(x, r)\} \leq \pi^{-d}\int_I |f(z(t))|\, dt,$$

where $z(t) \in D(x, 2r)$.

Proof of Theorem 3. The proof follows closely the one-type result in Biggins (1992). Let $x \in \Lambda_-$ be fixed. There is a $\alpha \in (1, 2]$, and a $B = B(x, 3r) \subset \Lambda_\alpha$, which means that $B \subset \Omega^2_\alpha$ and $B \subset \Omega^3_\alpha$. For any two positive integers n and N $(\geq n)$, we apply Lemma 2.4 to $\mathcal{W}_i^{N+1}(\lambda) - \mathcal{W}_i^n(\lambda)$ to get

$$\begin{aligned}\pi^d \sup_{\lambda \in B(x,r)} \left|\mathcal{W}_i^{N+1}(\lambda) - \mathcal{W}_i^n(\lambda)\right| &\leq \int_I \sum_{m=n}^{N} \left|\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))\right| dt, \\ &\leq \sum_{m=n}^{\infty} \int_I \left|\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))\right| dt,\end{aligned}$$

where $z(t) \in D(x, 2r) \subset B$. Applying Fubini's theorem and Lemma 2.3(iii), we get

$$E \sum_{m=n}^{\infty} \int_I \left|\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))\right| dt \leq \int_I \left[C\varphi^{1/\alpha} \frac{\kappa^{n/\alpha}}{1-\kappa^{1/\alpha}}\right] dt,$$

where the argument $z(t)$ has been suppressed on the right. Recall that $\kappa < 1$ throughout Ω^2_α and $\beta < \infty$ throughout Ω^3_α. Therefore, using the continuity of the various functions, there are constants $\delta < 1$ and $K < \infty$ such that the integrand on the right can be bounded by $K\delta^n$ throughout B and hence $D(x, 2r)$. Thus

$$E \sum_{m=n}^{\infty} \int_I \left|\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))\right| dt \leq K\delta^n \int_I dt < \infty$$

and so

$$\sum_{m=0}^{\infty} \int_I |\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))| dt < \infty, \quad \text{almost surely.}$$

Therefore $\{\mathcal{W}_i^n(\lambda)\}$ is a Cauchy sequence with supremum norm $||.||$ on $B(x, r)$, which implies uniform convergence on $B(x, r)$, almost surely.

For the remainder of this proof, the supremum norm is defined over $B(x, r)$. From the almost sure uniform convergence of the martingale, for fixed n, as $N \to \infty$,

$$||\mathcal{W}_i^N(\lambda) - \mathcal{W}_i^n(\lambda)|| \longrightarrow ||\mathcal{W}_i(\lambda) - \mathcal{W}_i^n(\lambda)||,$$

almost surely. Let n be fixed, then by taking expectation of both sides and applying Fatou's lemma, and then Lemma 2.4, we get

$$\begin{aligned}E||\mathcal{W}_i(\lambda) - \mathcal{W}_i^n(\lambda)|| &\leq \liminf_{N\to\infty} E||\mathcal{W}_i^N(\lambda) - \mathcal{W}_i^n(\lambda)|| \\ &\leq \pi^{-d} E \int_I \sum_{m=n}^{\infty} \left|\mathcal{W}_i^{m+1}(z(t)) - \mathcal{W}_i^m(z(t))\right| dt \\ &\leq \pi^{-d} K\delta^n \int_I dt < \infty \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

Finally, define

$$\Gamma = \bigcup_{x \in \Lambda_-} B(x, r).$$

Then, by a compactness argument, the uniform convergence of $\{\mathcal{W}_i^n(\lambda)\}$ in the open balls in Γ implies the uniform convergence in any compact subset of Γ. □

3 Proofs of Theorems 4 and 5

Let $1 \leq l \leq n$, and i, j be fixed. We start with a decomposition which is similar to (2.2),

$$\mathcal{W}_{ij}^n(\lambda) = \sum_k \sum_s \frac{v_k(\lambda)}{v_i(\lambda)} \frac{1}{\rho(\lambda)^l} \exp(-\lambda' z_{ik;s}^l) \mathcal{W}_{kj}^{n-l}(\lambda|l,s), \tag{3.1}$$

where, given $\mathcal{F}_l$, $\mathcal{W}_{kj}^{n-l}(\lambda|l,s)$ are independent copies of $\mathcal{W}_{kj}^{n-l}(\lambda)$ as s varies. Hence

$$E\left[\mathcal{W}_{ij}^n(\lambda)\middle|\,\mathcal{F}_l\right] = \frac{v_j(\lambda)}{v_i(\lambda)\rho(\lambda)^n} \sum_k \sum_s \exp(-\lambda' z_{ik;s}^l) m_{kj}^{n-l}(\lambda). \tag{3.2}$$

Therefore, given any sample path of the process, $E[\mathcal{W}_{ij}^n(\lambda)|\mathcal{F}_l]$ is an analytic function in $\lambda \in \Omega$.

For all $\lambda \in \Omega$ define

$$g_{ij}^n(\lambda \mid l) = \mathcal{W}_{ij}^n(\lambda) - E[\mathcal{W}_{ij}^n(\lambda) \mid \mathcal{F}_l],$$

and

$$h_{ij}^n(\lambda|l) = E[\mathcal{W}_{ij}^n(\lambda)|\mathcal{F}_l] - u_j(\lambda)v_j(\lambda)\mathcal{W}_i^l(\lambda).$$

The functions $g_{ij}^n(\lambda)$ and $h_{ij}^n(\lambda)$ are analytic functions in $\lambda \in \Omega$. Furthermore,

$$\left|\mathcal{W}_{ij}^n(\lambda) - u_j(\lambda)v_j(\lambda)\mathcal{W}_i(\lambda)\right| \leq \left|g_{ij}^n(\lambda|l)\right| + \left|h_{ij}^n(\lambda|l)\right| + |u_j(\lambda)v_j(\lambda)| \left|\mathcal{W}_i^l(\lambda) - \mathcal{W}_i(\lambda)\right|.$$

Hence $\left|\mathcal{W}_{ij}^n(\lambda) - u_j(\lambda)v_j(\lambda)\mathcal{W}_i(\lambda)\right|^\alpha$ is less than or equal to

$$3\left(\left|g_{ij}^n(\lambda|l)\right|^\alpha + \left|h_{ij}^n(\lambda|l)\right|^\alpha + |u_j(\lambda)v_j(\lambda)|^\alpha \left|\mathcal{W}_i^l(\lambda) - \mathcal{W}_i(\lambda)\right|^\alpha\right). \tag{3.3}$$

The idea now is to let $l = [n/2]$, where $[x]$ is the greatest integer not exceeding x, and n tend to infinity. This motivates the lemmas we now give before returning to the main proof. It is in these results that we need the uniformity proved in Theorem 1(iii).

Lemma 3.1 *Suppose $\vartheta \in (\Lambda_\alpha)_-$. Then there are positive constants c_1, c_2 and c_3, depending on α, a constant $\gamma < 1$ and a neighbourhood B of ϑ such that, for all $\lambda \in B$, $n = 1, 2, \ldots$, and all i, j,*

(i) $\left|h_{ij}^n(\lambda|l)\right| \leq c_1 \gamma^{n-l} \nu \phi^l \mathcal{W}_i^l(\theta)$,

(ii) $E\sum_{l=0}^n \left|h_{ij}^n(\lambda|l)\right|^\alpha \leq c_2 \phi^{n\alpha}$,

(iii) $\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)) \leq c_3 \phi^{n\alpha}$.

Proof. Using (3.2) and the definition of $\mathcal{W}_i^l(\lambda)$, at (1.4), we have

$$h_{ij}^n(\lambda|l) = \frac{v_j(\lambda)}{v_i(\lambda)\rho(\lambda)^l} \sum_k \sum_s \exp(-\lambda' z_{ik;s}^l) \left[\rho(\lambda)^{-(n-l)} m_{kj}^{n-l}(\lambda) - u_j(\lambda)v_k(\lambda)\right]. \tag{3.4}$$

Therefore, by Theorem 1(iii), there are constants $c_1 > 0$, $\gamma \in (0,1)$, and a neighbourhood B of ϑ such that, for all $\lambda \in B$ and all n,

$$\begin{aligned}\left|h_{ij}^n(\lambda|l)\right| &\leq c_1 \nu \phi^l \gamma^{n-l} \sum_k \sum_s \left[\frac{v_k(\theta)}{v_i(\theta)} \rho(\theta)^{-l} \exp(-\theta' z_{ik;s}^l)\right] \\ &= c_1 \nu \phi^l \gamma^{n-l} \mathcal{W}_i^l(\theta),\end{aligned} \tag{3.5}$$

as required. This and Lemma 2.3(iv) combine, after making B smaller if necessary, to show that, for some constant c_1', $E|h_{ij}^n(\lambda|l)|^\alpha \le c_1'\nu^\alpha(\gamma/\phi)^{(n-l)\alpha}\phi^{n\alpha}$. Since $\gamma < 1$ and $\phi \ge 1$, $(\gamma/\phi) < 1$, for all $\lambda \in B$. Thus

$$E\sum_{l=0}^{n}\left|h_{ij}^n(\lambda|l)\right|^\alpha \le c_1'\nu^\alpha\frac{1}{1-(\gamma/\phi)^\alpha}\phi^{n\alpha}.$$

The continuity of the functions ν and ϕ now implies (ii). Finally, by Lemma 2.2,

$$\begin{aligned}\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)) &\le C\sum_{l=1}^{n}E\left|E\left[\mathcal{W}_{ij}^n(\lambda)\middle|\mathcal{F}_l\right]-E\left[\mathcal{W}_{ij}^n(\lambda)\middle|\mathcal{F}_{l-1}\right]\right|^\alpha\\ &\le C'\left(|u_j(\lambda)v_j(\lambda)|^\alpha\sum_{l=1}^{n}E\left|\mathcal{W}_i^l(\lambda)-\mathcal{W}_i^{l-1}(\lambda)\right|^\alpha+2E\sum_{l=0}^{n}\left|h_{ij}^n(\lambda|l)\right|^\alpha\right).\end{aligned}$$

Thus, for δ small enough, by suitable bounding of the continuous functions involved in Lemma 2.3(i), for all $\lambda \in B(\vartheta,\delta)$

$$\sum_{l=1}^{\infty}E\left|\mathcal{W}_i^l(\lambda)-\mathcal{W}_i^{l-1}(\lambda)\right|^\alpha \le K_1 < \infty.$$

Then, bounding the continuous functions $u_j(\lambda)v_j(\lambda)$ and using part (ii) and the fact that $\phi \ge 1$ proves (iii). □

Lemma 3.2 *Suppose $\alpha \in (1,2]$ and $\vartheta \in (\Lambda_\alpha)_-$. Then there is a $B = B(\vartheta,\delta) \subset \Lambda_\alpha$, such that, for $l = [n/2]$,*

(i) for all $\lambda \in B$, $h_{ij}^n(\lambda|l) \longrightarrow 0$ as $n \to \infty$, almost surely and in αth mean;

(ii) $\sup\left\{|h_{ij}^n(\lambda|l)| : \lambda \in B\right\} \longrightarrow 0$ as $n \to \infty$, almost surely and in mean.

Proof. Let $B = B(\vartheta,\delta)$ be a neighbourhood of ϑ over which (3.5) holds. Take δ smaller if necessary, so that $\phi\gamma < 1$ throughout B. Applying Theorems 2 and 3 to $\mathcal{W}_i^l(\theta)$ and simple bounding now give all the claimed results. □

Lemma 3.3 *Suppose $\alpha \in (1,2]$ and $\vartheta \in (\Lambda_\alpha)_-$. Then there is a neighbourhood of ϑ, say $B = B(\vartheta,r) \subset \Lambda_\alpha$, such that, for $l = [n/2]$,*

(i) for all $\lambda \in B$, $g_{ij}^n(\lambda|l) \longrightarrow 0$ as $n \to \infty$, almost surely and in αth mean;

(ii) $\sup\left\{|g_{ij}^n(\lambda|l)| : \lambda \in B\right\} \longrightarrow 0$ as $n \to \infty$, almost surely and in mean.

Proof. Note first that, using (3.1) and Lemma 2.2,

$$\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)\big|\,\mathcal{F}_l) \le C\bar{v}(\lambda)^\alpha \mid \rho(\lambda)\mid^{-l\alpha}\sum_k\sum_s \exp(-\alpha\theta' z_{ik;s}^l)\sigma^\alpha(\mathcal{W}_{kj}^{n-l}(\lambda)). \tag{3.6}$$

Let $B = B(\vartheta,\delta)$ be inside the neighbourhood of ϑ introduced in Lemma 3.1. Then for some positive constant c and for all $\lambda \in B$, $n = 1,\ldots,$ and all i,j, applying Lemma 3.1(iii) to the bound (3.6) gives

$$\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)\big|\,\mathcal{F}_l) \le c\nu_1\mathcal{W}_i^l(\alpha\theta)\phi^{\alpha(n-l)}\kappa^l. \tag{3.7}$$

Now, since $E[\mathcal{W}_i^l(\alpha\theta)] = 1$,

$$E\left[\left|g_{ij}^n(\lambda|l)\right|^\alpha\right] = E\left[\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)\middle|\,\mathcal{F}_l)\right] \le c\nu_1\phi^{\alpha(n-l)}\kappa^l.$$

Choose δ smaller if necessary so that $\phi^\alpha\kappa < 1$ for all $\lambda \in B$. Then the right hand side of the above inequality converges to zero geometrically quickly, as $n \to \infty$, when $l = [n/2]$. This implies the convergence of $g_{ij}^n(\lambda|l) \to 0$, in αth mean and almost surely.

Jensen's inequality and (3.7) give

$$E\left|g_{ij}^n(\lambda|l)\right| \le E\left[\sigma^\alpha(\mathcal{W}_{ij}^n(\lambda)|\mathcal{F}_l)\right]^{1/\alpha} \le c^{1/\alpha}\nu_1^{1/\alpha}\phi^{n-l}\kappa^{l/\alpha}.$$

For $l = [n/2]$, $\phi^{n-l}\kappa^{l/\alpha}$ is asymptotically equivalent to $(\phi\kappa^{1/\alpha})^l$, and $\phi\kappa^{1/\alpha} \le \gamma_1 < 1$ in B. Hence there is a $C < \infty$ such that for all $\lambda \in B$, $n = 1, 2, \ldots$, $l = [n/2]$, and all i, j, $E\left|g_{ij}^n(\lambda|l)\right| \le C\gamma_1^l$. Let $2r < \delta$ and apply Lemma 2.4 to the analytic functions $g_{ij}^n(\lambda|l)$ to get

$$\sup_{\lambda\in B(\vartheta,r)}\left|g_{ij}^n(\lambda|l)\right| \le \pi^{-d}\int_I\left|g_{ij}^n(z(t)|l)\right|dt,$$

where $z(t) \in D(\vartheta, 2r)$. The expectation of the right hand side here goes to zero geometrically quickly. Hence $g_{ij}^n(\lambda|l)$ converges to zero uniformly in $\lambda \in B(\vartheta, r)$ almost surely and in mean. □

Proof of Theorems 4 and 5. In (3.3), let $l = [n/2]$ and then let $n \to \infty$. In a suitable neighbourhood of $\vartheta \in (\Lambda_\alpha)_-$, the almost sure and αth mean convergence of $\{\mathcal{W}_i^n(\lambda) - \mathcal{W}_i(\lambda)\}$, $g_{ij}^n(\lambda|l)$ and $h_{ij}^n(\lambda|l)$ are contained in Theorem 2, Lemma 3.3(i) and Lemma 3.2(i) respectively. Now take the union of these neighbourhoods as Γ_1. This proves Theorem 4.

Similarly, the almost sure and mean uniform convergence of $\{\mathcal{W}_i^n(\lambda) - \mathcal{W}_i(\lambda)\}$, $g_{ij}^n(\lambda|l)$ and $h_{ij}^n(\lambda|l)$ are contained in Theorem 3, Lemma 3.3(ii) and Lemma 3.2(ii) respectively. A union of suitable neighbourhoods now provides Γ_2, proving Theorem 5. □

4 Proof of Theorem 6

Lemma 4.1 *If the branching random walk is strongly non-lattice then for any compact set $\mathcal{K} \subset \{\theta \in \mathbb{R}^d : \sup_{i,j} m_{ij}(\theta) < \infty\}$ and all $a \in (0,1)$ there is an $\varepsilon < 1$ such that*

$$\varepsilon^{-n}\sup_{i,j}\sup_{\theta\in\mathcal{K}}\sup_{a\le|\eta|\le a^{-1}}\left|\frac{m_{ij}^n(\theta+\mathbf{i}\eta)}{\rho^n(\theta)}\right| \to 0 \quad \text{as } n \to \infty. \tag{4.8}$$

Proof. Let (k, l) be such that (1.7) holds. As a function of η, $m_{kl}(\theta + \mathbf{i}\eta)/m_{kl}(\theta)$ is the characteristic function of the probability measure $m_{kl}(\theta)^{-1}e^{-\theta' x}\mu_{kl}(dx)$ and (1.7) implies that this measure is strongly non-lattice. Then, using this and continuity,

$$\tilde{\varepsilon} = \sup_{\theta\in\mathcal{K}}\sup_{a\le|\eta|\le a^{-1}}\left|\frac{m_{kl}(\theta+\mathbf{i}\eta)}{m_{kl}(\theta)}\right| < 1.$$

Let $M^\dagger(\theta)$ be the matrix with (k, l)th entry $\tilde{\varepsilon}m_{kl}(\theta)$ and with other entries the same as those in $M(\theta)$ and let $(M^\dagger(\theta)^n)_{ij}$ be the (i, j)th entry of its nth power. Let $\rho^\dagger(\theta)$

be the eigenvalue of $M^\dagger(\theta)$ of maximum modulus, with right eigenvector $v^\dagger(\theta)$. Then $\rho^\dagger(\theta) < \rho(\theta)$ (see Seneta (1973, Theorem 1.1)). Furthermore, since Theorem 1 applies to $M^\dagger(\theta)$, there is an $\varepsilon < 1$ such that $\rho^\dagger(\theta) < \varepsilon\rho(\theta)$ for all $\theta \in \mathcal{K}$; also,

$$\sup_{\theta\in\mathcal{K}} \max_{i,j}\{v_i^\dagger(\theta)/v_j^\dagger(\theta)\} = c < \infty.$$

Now,

$$\sup_{a\le|\eta|\le a^{-1}} |(M(\theta+\mathbf{i}\eta)^n)_{ij}| \le (M^\dagger(\theta)^n)_{ij} \le c(\varepsilon\rho(\theta))^n,$$

proving (4.8). □

Proof of Theorem 6.

It is sufficient to prove (1.8) with the compact set $\mathcal{K}$ some closed ball in $\mathbb{R}^d$ centred on $\vartheta \in \Lambda_-$, since a simple covering argument extends the result to general $\mathcal{K}$. The proof starts by introducing various bounds which lead to the appropriate ball to consider. Recall that $S(x,r)$ is a ball in $\mathbb{R}^d$ centred on x with radius r and $\overline{S}(x,r)$ is its closure.

Choose $\vartheta \in \Lambda_-$, then $\vartheta \in \Lambda_\alpha$ for some α, which is now fixed. Now take δ_1 such that $S(\vartheta, 2\delta_1) \in \Lambda_\alpha$ and $S_1 = \overline{S}(\vartheta,\delta_1)$. Let ε_1 be an ε such that (4.8) holds when $\mathcal{K} = \{\theta \in S_1\}$ and $a = \epsilon/2$. Recall that $\rho(\alpha\theta)^{1/\alpha}/\rho(\theta) < 1$ for $\theta \in \Lambda_\alpha$. Let

$$\varepsilon_2 = \max\left\{\varepsilon_1, \sup_{\theta\in S_1} \frac{\rho(\alpha\theta)^{1/\alpha}}{\rho(\theta)}\right\}.$$

Now let $S_2 = \overline{S}(\vartheta, 2\delta_2)$ with $2\delta_2 < \delta_1$ and let

$$\overline{\rho} = \sup_{\theta\in S_2} \rho(\theta) \quad \text{and} \quad \underline{\rho} = \inf_{\theta\in S_2} \rho(\theta).$$

Now take δ_2 small enough that

$$\varepsilon_3 = \varepsilon_2 \frac{\sup_{\theta\in S_2} \rho(\theta)}{\inf_{\theta\in S_2} \rho(\theta)} = \varepsilon_2 \frac{\overline{\rho}}{\underline{\rho}} < 1. \tag{4.9}$$

Let $S = \overline{S}(\vartheta, \delta_2) \subset S_2 = \overline{S}(\vartheta, 2\delta_2)$. Define the two regions

$$R = \{\theta + \mathbf{i}\eta : \theta \in S, \epsilon \le \eta \le \epsilon^{-1}\} \quad \text{and} \quad R_2 = \{\theta + \mathbf{i}\eta : \theta \in S_2, \epsilon/2 \le \eta \le 2\epsilon^{-1}\}.$$

We will prove (1.8) for $\theta \in S$, that is for $\lambda \in R$. Clearly

$$\sup_{\lambda\in R} |\mathcal{M}_{ij}^n(\lambda)| \le \sup_{\lambda\in R} |\mathcal{M}_{ij}^n(\lambda) - m_{ij}^n(\lambda)| + \sup_{\lambda\in R} |m_{ij}^n(\lambda)|. \tag{4.10}$$

Furthermore,

$$\sup_{\lambda\in R} |m_{ij}^n(\lambda)| \le \sup_{\lambda\in R_2} |m_{ij}^n(\lambda)| = \sup_{\lambda\in R_2} \left|\frac{m_{ij}^n(\lambda)}{\rho(\theta)^n}\rho(\theta)^n\right| \le (\varepsilon_1\overline{\rho})^n \le (\varepsilon_2\overline{\rho})^n. \tag{4.11}$$

Using Lemma 2.2,

$$E\left|\mathcal{M}_{ij}^n(\lambda) - m_{ij}^n(\lambda)\right|^\alpha \le C\sum_{l=0}^{n-1} E\left|E[\mathcal{M}_{ij}^n(\lambda)|\mathcal{F}_{l+1}] - E[\mathcal{M}_{ij}^n(\lambda)|\mathcal{F}_l]\right|^\alpha \tag{4.12}$$

and splitting on the lth generation shows that

$$\begin{aligned} E[\mathcal{M}^n_{ij}(\lambda)|\mathcal{F}_l] &= \sum_k \left(\sum_s \exp(-\lambda' z^l_{ik;s})\right) m^{n-l}_{kj}(\lambda) \\ &= \sum_k \left(\int \exp(-\lambda' z) Z^l_{ik}(dz)\right) m^{n-l}_{kj}(\lambda). \end{aligned} \tag{4.13}$$

Following (2.1), let

$$\mathcal{M}^{n-l}_{kj}(\lambda|l,s) = \int_{\mathbb{R}^d} e^{-\lambda' x} Z^n_{kj}(dx|l,s),$$

then

$$\int \exp(-\lambda' z) Z^{l+1}_{ik}(dz) = \sum_h \sum_s \exp(-\lambda' z^l_{ih;s}) \mathcal{M}^1_{hk}(\lambda|l,s). \tag{4.14}$$

Let $\Delta_l = E[\mathcal{M}^n_{ij}(\lambda)|\mathcal{F}_{l+1}] - E[\mathcal{M}^n_{ij}(\lambda)|\mathcal{F}_l]$. Then, using (4.13) and (4.14)

$$\Delta_l = \sum_k \left(\sum_h \sum_s \exp(-\lambda' z^l_{ih;s})(\mathcal{M}^1_{hk}(\lambda|l,s) - m_{hk}(\lambda))\right) m^{n-l-1}_{kj}(\lambda)$$

and, given $\mathcal{F}_l$, $\{\mathcal{M}^1_{hk}(\lambda|l,s) - m_{hk}(\lambda) : s\}$ are independent variables with zero mean. We need a bound on the moments of these variables. Let

$$c_1 = 4\left(\sup_{i,j} \sup_{\theta \in S_2} E\left|\mathcal{M}^1_{ij}(\theta)\right|^\alpha\right)^{1/\alpha},$$

which is finite because $S_2 \subset \Lambda_\alpha$. Then simple bounding gives

$$\sup\left\{E\left|\mathcal{M}^1_{ij}(\lambda) - m_{ij}(\lambda)\right|^\alpha : \lambda = \theta + \mathbf{i}\eta, \theta \in S_2, i, j\right\} \le (c_1)^\alpha.$$

Now, applying Lemma 2.2 again and the various bounds, for any $\lambda \in R_2$

$$\begin{aligned} E|\Delta_{l-1}|^\alpha &\le C \sum_k \left(\sum_h m^{l-1}_{ih}(\alpha\theta) E\left|\mathcal{M}^1_{hk}(\lambda) - m_{hk}(\lambda)\right|^\alpha\right) |m^{n-l}_{kj}(\lambda)|^\alpha \\ &\le c_2(c_1)^\alpha \rho(\alpha\theta)^{l-1} (\varepsilon_1 \overline{\rho})^{(n-l)\alpha} \\ &\le c_2(c_1)^\alpha (\varepsilon_2 \overline{\rho})^{(l-1)\alpha} (\varepsilon_1 \overline{\rho})^{(n-l)\alpha} \\ &\le (c_3 c_1 (\varepsilon_2 \overline{\rho})^n)^\alpha, \end{aligned}$$

where c_1 and c_3 are independent of $\theta \in S_2$. Hence (4.12) gives

$$E\left|\mathcal{M}^n_{ij}(\lambda) - m^n_{ij}(\lambda)\right|^\alpha \le n\,(c_3 c_1 (\varepsilon_2 \overline{\rho})^n)^\alpha$$

for any $\lambda \in R_2$.

The region R can be covered with a finite number of polydiscs such that when their radii are doubled they still lie inside R_2. Hence, using (4.12), the bound just obtained and Lemma 2.4, for a suitable constant c_4,

$$E \sup_{\lambda \in R} |\mathcal{M}^n_{ij}(\lambda) - m^n_{ij}(\lambda)| \le c_4 (\varepsilon_2 \overline{\rho})^n n^{1/\alpha}.$$

Combining this with (4.10) and (4.11)

$$E \sup_{\lambda \in R} |\mathcal{M}_{ij}^n(\lambda)| \leq (c_4 n^{1/\alpha} + c_1)(\varepsilon_2 \overline{\rho})^n$$

and so, using (4.9),

$$E \sup_{\lambda \in R} \left| \frac{\mathcal{M}_{ij}^n(\lambda)}{\rho(\theta)^n} \right| \leq \frac{1}{\underline{\rho}^n} E \sup_{\lambda \in R} \left| \mathcal{M}_{ij}^n(\lambda) \right| \leq (c_4 n^{1/\alpha} + c_1) \varepsilon_3^n.$$

Hence (1.8) holds for any $\varepsilon \in (\varepsilon_3, 1)$. □

5 Approximation of measures

This section contains the preparatory work on saddlepoint approximations for the proof of Theorem 7. The idea is to develop explicit estimates that apply to a particular measure through only a few of its attributes. Most treatments of these matters, with the application to variables that are, or look like, sums of independent identically distributed random variables in mind, bring n into the picture sooner than we do. The treatment draws on ideas from Stone (1967), von Bahr (1967), Bhattacharya (1972, 1977), Chaganty and Sethuraman (1993), and Jensen (1995).

Recall that $S(x, \epsilon) = \{y : |x - y| < \epsilon\}$. Let Γ be a probability measure with $\Gamma(S(0, 1)) > 1/2$ and characteristic function $\int e^{\mathbf{i}tu}\Gamma(du)$ vanishing outside $\{t : |t| \geq \zeta\}$. Let Γ_ϵ be the measure given by $\Gamma_\epsilon(A) = \Gamma(\{x : \epsilon x \in A\})$, so that its characteristic function is zero outside $\{t : |t| \geq \epsilon^{-1}\zeta\}$. Let ν and μ be probability measures, b_ϵ be the supremum of the modulus of the density of $(\nu - \mu) * \Gamma_\epsilon$ and q be the supremum of the density of μ. Let $\mathfrak{C}$ be the convex sets in $\mathbb{R}^d$.

The set $\mathfrak{G}'(G)$ is made up of those $f \in \mathfrak{D}$ for which there is a bounded decreasing function g with $\int g(r)(r + 1)^{d-1} dr = G$ and

$$\sup\{|f(x)|, |f'(x)| : |x| \geq (r - 1)\} \leq g(r).$$

Note that $\mathfrak{G}(G)$ is defined in the same way, except that $\int g(r)(r+1)^{d+1} dr = G$; therefore, $\mathfrak{G}(G) \subset \mathfrak{G}'(G)$.

Lemma 5.1 *For a constant Δ, depending only on the dimension, and $\epsilon < 1$*

$$\sup\left\{\left|\int_C f(x + u) d(\nu - \mu)(x)\right| : u \in \mathbb{R}^d, C \in \mathfrak{C}, f \in \mathfrak{G}'(G)\right\} \quad \leq \quad G\Delta\left[b_\epsilon + \epsilon q\right].$$

Proof. Let

$$\omega_f(x, \epsilon) = \sup\{|f(y) - f(z)| : y, z \in S(x, \epsilon)\}$$

and

$$\overline{f}^{\,\epsilon}(x) = \sup\{|f(y)| : y \in S(x, \epsilon)\}.$$

Let $\gamma \equiv \Gamma(S(0, 1))$. Then, applying Lemma 2.2 in Bhattacharya (1972) and the remark following it,

$$\sup_u \left|\int f(x + u) d(\nu - \mu)(x)\right| \quad \leq \quad \frac{1}{2\gamma - 1}\left[b_\epsilon \int \overline{f}^{\,\epsilon}(x) dx + q \int \omega_f(x, 2\epsilon) dx\right] \quad (5.15)$$

For a set $A \subset \mathbb{R}^d$, let $A^\epsilon = \{y : |y - A| < \epsilon\}$ and ∂A be the boundary of A. Let $\mathfrak{S}$ be the surface area of the unit ball. Then, for $C \in \mathfrak{C}$, from (15) in von Bahr (1967),

$$\int I(x \in (\partial C)^\epsilon) g(|x|) dx \leq 2\mathfrak{S}G\epsilon.$$

Take $\epsilon < 1$. Let $h(x) = f(x)I(x \in C)$. If $x \in (\partial C)^\epsilon$ then $\omega_h(x,\epsilon) \leq 2\sup\{|f(y)| : y \in S(x,\epsilon)\} \leq 2g(|x|)$. If $x \in S(x,\epsilon) \subset C$ then, using Taylor's theorem, $\omega_h(x,\epsilon) = \omega_f(x,\epsilon) \leq 2\epsilon g(|x|)$. Hence

$$\omega_h(x,\epsilon) \leq 2g(|x|)\left[I(x \in (\partial C)^\epsilon) + \epsilon I(x \in C)\right]$$

and so

$$\int \omega_h(x,\epsilon)dx \leq 2\int \left[I(x \in (\partial C)^\epsilon) + \epsilon I(x \in C)\right] g(|x|)dx \leq 6\epsilon\mathfrak{S}G.$$

Also

$$\int \overline{h}^{\,\epsilon}(x)dx \leq \int \overline{f}^{\,\epsilon}(x)dx \leq \int g(|x|)dx = \mathfrak{S}G.$$

Substituting these into (5.15) gives the result. □

Recall that for $A \subset \mathbb{C}^d$, A_- is its intersection with $\mathbb{R}^d$. Let $\mathcal{K}$ be compact in $\mathbb{R}^d$, $\tau < 1$, and let

$$\mathcal{K}_\tau = \{\theta + \lambda \in \mathbb{C}^d : \theta \in \mathcal{K}, |\lambda| \leq \tau\}.$$

and let B be an open set containing $\mathcal{K}_\tau$ with B_- convex. For Z a measure on $\mathbb{R}^d$ let

$$\widehat{Z}(\theta) = \int e^{-\theta' x} Z(dx).$$

Suppose that $\log \widehat{Z}$ is analytic on B and strictly convex on B_-. Let $-m(\theta)$ be the vector of its first derivatives and $\Sigma(\theta)$ be the matrix of its second derivatives on B_-. Then $\Sigma(\theta)$ is positive-definite because $\log \widehat{Z}$ is strictly convex. Let $\mathfrak{u}$ be a (finite) bound on the modulus of all derivatives of $\log \widehat{Z}$ up to and including order 3 over $\mathcal{K}_\tau$. Let $\mathfrak{c}$ be a lower bound on the smallest eigenvalue of $\Sigma(\theta)$ as θ varies through $\mathcal{K}$ and let $\mathfrak{r} = \mathfrak{u}/\mathfrak{c}$. Since $\log \widehat{Z}$ is strictly convex and analytic on B_-, we can take $\mathfrak{c} > 0$. Also, let

$$\varepsilon(a) = a^{-d} \sup\left\{ \left| \frac{\widehat{Z}(\theta + \mathbf{i}t)}{\widehat{Z}(\theta)} \right| : \theta \in \mathcal{K},\ a \leq |t| < 1/a \right\}.$$

Fix $\vartheta \in \mathcal{K}$. Let ν be the probability measure given by

$$\nu(A) = \frac{\int I_A(x) e^{-\vartheta' x} Z(dx)}{\int e^{-\vartheta' y} Z(dy)},$$

with mean m $(= m(\vartheta))$ and covariance matrix Σ $(= \Sigma(\vartheta))$. Let $\widehat{\nu}(\lambda)$ be the Laplace transform of ν, so that $\widehat{\nu}(\lambda) = \widehat{Z}(\vartheta + \lambda)/\widehat{Z}(\vartheta)$. Let μ be the normal distribution with mean m and covariance matrix Σ. The idea is to use μ to approximate ν, and hence Z, through Lemma 5.1, in a way that is suitably uniform. Note that ν and μ both depend on ϑ; to emphasise this in the next definition we use ν_ϑ and μ_ϑ. Let

$$A(Z) = \sup\left\{ \left| \int_C f(x+u)(\nu_\vartheta - \mu_\vartheta)(dx) \right| : u \in \mathbb{R}^d, f \in \mathfrak{G}'(G), C \in \mathfrak{C}, \vartheta \in \mathcal{K} \right\}.$$

Obviously by bounding $A(Z)$ we allow f to be shifted arbitrarily, f to vary within $\mathfrak{G}'(G)$ and integration over an arbitrary convex set.

Theorem 9 *For a constant Δ', depending only on the dimension, $\delta \leq \min\left\{\tau, e^{-d}/(4\mathfrak{r})\right\}$ and $\epsilon \leq \min\{1, \zeta/\delta, \zeta\delta\}$,*

$$A(Z) \leq G\Delta'\left[\frac{\mathfrak{r}}{\mathfrak{c}^{(d+1)/2}} + \varepsilon\left(\frac{\epsilon}{\zeta}\right) + \frac{(e^{-\delta^2\mathfrak{c}/4} + \epsilon)}{\mathfrak{c}^{d/2}}\right].$$

Proof. Let q be a bound on the density of μ and let b_ϵ be a bound on the modulus of the density of $(\nu - \mu) * \Gamma_\epsilon$. Note first that $q \leq (2\pi\mathfrak{c})^{-d/2}$, which is independent of ϑ.

For $|\lambda| \leq \tau$, let

$$\begin{aligned}\psi(\lambda) &= \log\widehat{\nu}(\lambda) - m'\lambda + \frac{1}{2}\lambda'\Sigma\lambda \\ &= \log\widehat{Z}(\vartheta + \lambda) - \log\widehat{Z}(\vartheta) - m'\lambda + \frac{1}{2}\lambda'\Sigma\lambda\end{aligned}$$

Then, by arrangement, ψ is analytic on $|\lambda| \leq \tau$ with $\psi(0) = 0$ and all its first and second derivatives vanishing at 0.

Using Taylor's theorem and the analyticity of $\log\widehat{Z}$ on B, for $|\lambda| \leq \tau$,

$$|\psi(\lambda)| \leq \frac{d^3}{3!}\mathfrak{u}|\lambda|^3 \leq e^d\mathfrak{u}|\lambda|^3$$

and then, for $|\lambda| \leq \delta$,

$$|\psi(\lambda)| \leq e^d\mathfrak{u}|\lambda|^3 \leq \left(4e^d\frac{\mathfrak{u}}{\mathfrak{c}}\delta\right)\frac{\mathfrak{c}}{4}|\lambda|^2 \leq \frac{\mathfrak{c}}{4}|\lambda|^2.$$

Using these two inequalities gives

$$|\exp\{\psi(\lambda)\} - 1| \leq e^d\mathfrak{u}|\lambda|^3\exp\left\{\frac{\mathfrak{c}}{4}|\lambda|^2\right\}.$$

The key point is that right hand side here does not depend on ϑ. Using this bound,

$$\begin{aligned}\int_{|t|<\delta}\left|\widehat{\nu}(t) - e^{\mathbf{i}m't}e^{-t'\Sigma t/2}\right|dt &\leq \int_{|t|<\delta} e^{-t'\Sigma t/2}\left|e^{\psi(t)} - 1\right|dt \\ &\leq e^d\mathfrak{u}\int_{|t|<\delta}|t|^3e^{\mathfrak{c}|t|^2/4}e^{-t'\Sigma t/2}dt \\ &\leq e^d\mathfrak{u}\int_{|t|<\delta}|t|^3e^{-|t|^2\mathfrak{c}/4}dt \\ &\leq e^d\mathfrak{u}\frac{1}{\mathfrak{c}^{(d+3)/2}}\int|z|^3\,e^{-|z|^2/4}dz \\ &\leq \frac{\Delta_1\mathfrak{r}}{\mathfrak{c}^{(d+1)/2}},\end{aligned}$$

where Δ_1 depends only on the dimension, d. Then

$$\begin{aligned}(2\pi)^d b_\epsilon &\leq \int_{|t|<\zeta/\epsilon}\left|\widehat{\nu}(t) - e^{\mathbf{i}m't}e^{-t'\Sigma t/2}\right|dt \\ &\leq \int_{|t|<\delta}\left|\widehat{\nu}(t) - e^{\mathbf{i}m't}e^{-t'\Sigma t/2}\right|dt + \int_{\delta\leq|t|<\zeta/\epsilon}|\widehat{\nu}(t)|\,dt + \int_{\delta\leq|t|}e^{-t'\Sigma t/2}dt \\ &\leq \frac{\Delta_1\mathfrak{r}}{\mathfrak{c}^{(d+1)/2}} + \varepsilon\left(\frac{\epsilon}{\zeta}\right)\frac{\mathfrak{S}}{d} + e^{-\delta^2\mathfrak{c}/4}\frac{(4\pi)^{d/2}}{\mathfrak{c}^{d/2}}.\end{aligned}$$

Substituting these estimates for b_ϵ and q into Lemma 5.1 gives the result. □

For the main result, we also need to approximate the normal distribution μ. To formulate the theorem, let

$$B(v) = \sup\left\{\left|\int_C f(x)\mu(dx+m-v) - \frac{\int_C f(x)dx}{\sqrt{(2\pi)^d \det[\Sigma]}}\right| : f \in \mathfrak{G}(G), C \in \mathfrak{C}, \vartheta \in \mathcal{K}\right\}.$$

Recall that μ, m and Σ each depend on ϑ.

Theorem 10 *For a constant Δ'' depending only on the dimension,*

$$B(v) \leq \Delta'' \frac{1+|v|^2}{\mathfrak{c}^{(d+2)/2}} G.$$

Proof. Temporarily, let $h(x) = f(x)I(x \in C)$ and, for a fixed $\vartheta \in \mathcal{K}$,

$$B_* = \left|\sqrt{(2\pi)^d \det[\Sigma]}\int h(x)\mu(dx+m-v) - \int h(x)dx\right|.$$

Then

$$\begin{aligned}
B_* &= \left|\int h(x+v)\left(\exp(-x'\Sigma^{-1}x/2) - 1\right)dx\right| \\
&\leq \frac{1}{2}\int |f(x+v)|(x'\Sigma^{-1}x)dx \\
&\leq \frac{1}{2\mathfrak{c}}\int |f(x)||x-v|^2 dx \\
&\leq \frac{1}{2\mathfrak{c}}\int g(|x|)|x|^2 dx + \frac{|v|^2}{2\mathfrak{c}}\int g(|x|)dx \\
&\leq \frac{1+|v|^2}{2\mathfrak{c}}\mathfrak{S}G,
\end{aligned}$$

and so

$$\frac{B_*}{\sqrt{(2\pi)^d \det[\Sigma]}} \leq \frac{B_*}{\sqrt{(2\pi)^d \mathfrak{c}^d}} \leq \frac{\mathfrak{S}}{2\sqrt{(2\pi)^d}}\frac{1+|v|^2}{\mathfrak{c}^{(d+2)/2}}G,$$

as required. □

It is worth, very briefly, relating these results to those in Chaganty and Sethuraman (1993). Temporarily following the notation used there, let T_n be a univariate random variable with moment generating function $\exp(n\psi_n(z))$. Suppose that ψ_n is analytic, and bounded in n and z, on Ω, and has a second derivative bounded below by α there. Let $\{\tau_n\}$ be a positive bounded sequence inside Ω_- and suppose that, for any $a > 0$,

$$\varepsilon_n(a) = a\sup\left\{\left|\frac{\exp(n\psi_n(\tau_n + \mathbf{i}t))}{\exp(n\psi_n(\tau_n))}\right| :\ a \leq |t| < 1/a\right\} = o(n^{-1/2}).$$

These are the conditions of Theorem 3.3 in Chaganty and Sethuraman (1993). Under these conditions we can take $\mathfrak{c} = n\alpha$, $\mathfrak{u} = nU$ and $\mathfrak{r}$ is bounded.

Take $f(x)$ to be zero on $(-\infty, -1]$, $e^{-\tau x}$ on $[0, \infty)$ and differentiable, with a bounded derivative, in between and at -1 and 0. For $\tau \in (0, B]$ it is routine calculus to show that there is a constant C such that f is in $\mathfrak{G}'(C/\tau)$ and $\mathfrak{G}(C/\tau^3)$. Apply Theorems 9 and 10 with this f, the convex set C being $(0,\infty)$ and $\mathcal{K}$ being the point set $\{\tau_n\}$ when estimating the nth distribution, to get

$$\frac{P\left[(T_n - n\psi_n'(\tau_n)) \in (0,\infty)\right]}{\exp(n\psi_n(\tau_n) - n\tau_n\psi'(\tau_n))} - \frac{1}{\tau_n\sqrt{2\pi n\psi_n''(\tau_n)}} = o\left(\frac{1}{\tau_n\sqrt{n}}\right),$$

which contains the conclusion of Theorem 3.3 in Chaganty and Sethuraman (1993). Clearly, a multivariate version of this result would also follow directly from the discussion here, as would a result uniform in τ in compact subsets of $(0,\infty)\cap\Omega_-$. Also, when sums of independent identically distributed variables are considered, so that $\psi_n(z)$ does not depend on n, the result contained in Theorem 3 in Stone (1967) is easily derived (for the non-lattice case).

6 Proof of Theorem 7 and its Corollary

Proof of Theorem 7.

Note first that $P(\mathcal{W}_i(\theta)=0)$ is a fixed point of the multivariate generating function of the underlying Galton-Watson process. Hence when the martingale converges in mean, so that $E\mathcal{W}_i(\theta)=1$, these probabilities must be less than one and so must equal the extinction probabilities from that starting type. Thus $\mathcal{W}_i(\theta)>0$ agrees with the survival set, almost surely for $\theta\in\Lambda_-$. The continuity of $\mathcal{W}_i(\theta)$ now means the null set can be chosen independent of θ. For the rest of the calculation we deal with sample paths in $\mathcal{S}$. Then there is an N such that $\mathcal{W}_{ij}^n(\theta)>0$ for all $n\geq N$. We take $n\geq N$.

Let $\mathcal{K}$ be a compact subset of Λ_- and fix j. Let the function w^n be defined by

$$w^n(\lambda)=\log\left(\frac{\mathcal{W}_{ij}^n(\lambda)v_i(\lambda)}{v_j(\lambda)}\right).$$

Using Theorems 1(ii), 3 and 5, there is an open B (with $\Lambda_-\subset B\subset\Lambda$) such that w^n is analytic in $\lambda\in B$ and, using Theorems 1(ii), 3 and 5, it converges to an analytic function there, and for some τ

$$\mathcal{K}_\tau=\{\theta+\lambda\in\mathbb{C}^d:\theta\in\mathcal{K},|\lambda|\leq\tau\}\subset B.$$

Analyticity of w^n and its limit on B mean that all its derivatives are uniformly bounded on $\mathcal{K}_\tau$.

To make the connection with the previous section, for fixed j and n let the measure Z be Z_{ij}^n and, for fixed θ, let $\nu(dx)=e^{-\theta' x}Z(dx)/\int e^{-\theta' y}Z(dy)$. Note first that, with $f(x)=e^{\theta' x}h(x)$ and $v_1=v_1(\theta)$

$$\begin{aligned}
e^{n\xi(\theta)}\int_C h(x)Z_{ij}^n(dx+nv_1(\theta)) &= \frac{e^{-n\theta v_1(\theta)}}{\rho(\theta)^n}\int_C f(x)e^{-\theta' x}Z_{ij}^n(dx+nv_1(\theta))\\
&= \frac{\int e^{-\theta x}Z_{ij}^n(dx)}{\rho(\theta)^n}\int_C f(x)\nu(dx+nv_1)\\
&= \frac{v_i(\theta)\mathcal{W}_{ij}^n(\theta)}{v_j(\theta)}\int_C f(x)\nu(dx+nv_1).
\end{aligned}$$

Now, by Theorem 5, $v_i(\theta)\mathcal{W}_{ij}^n(\theta)/v_j(\theta)$ converges to $v_i(\theta)u_j(\theta)\mathcal{W}_i(\theta)$ uniformly in $\theta \in \mathcal{K}$. By assumption, $f \in \mathfrak{G}(G)$ and so the results of the previous section can be applied to consider the convergence of $\int_C f(x)\nu(dx + n\upsilon_1)$ once we show Z satisfies the appropriate conditions.

For $\lambda \in B$, substitution shows that

$$\log\left(\widehat{Z}(\lambda)\right) - n\log(\rho(\lambda)) \quad = \quad w^n(\lambda).$$

It follows from the uniform boundedness of the derivatives of w^n that $\{m - n\upsilon_1 : n\}$ and $\{\Sigma - n\upsilon_2 : n\}$ are bounded uniformly in $\theta \in \mathcal{K}$. Furthermore there are constants $l > 0$ and $L < \infty$ such that, for all large enough n, $\mathfrak{c}/n \geq l$ and $\mathfrak{u}/n \leq L$ and then $\mathfrak{r} \leq L/l$.

For any $\epsilon > 0$,

$$\begin{aligned}
\varepsilon(\epsilon) &= \epsilon^{-d}\sup\left\{\left|\frac{\widehat{Z}(\theta + \mathbf{i}\eta)}{\widehat{Z}(\theta)}\right| : \theta \in \mathcal{K}, \epsilon \leq |\eta| < 1/\epsilon\right\} \\
&= \epsilon^{-d}\sup\left\{\frac{v_j(\theta)}{v_i(\theta)\mathcal{W}_{ij}^n(\theta)}\left|\frac{\mathcal{M}_{ij}^n(\theta + \mathbf{i}\eta)}{\rho(\theta)^n}\right| : \theta \in \mathcal{K}, \epsilon \leq |\eta| < 1/\epsilon\right\} \\
&\to 0
\end{aligned}$$

geometrically quickly as $n \to \infty$, using Theorems 5 and 6. Hence, using Theorem 9, $n^{d/2}A(Z_{ij}^n) \to 0$. Furthermore, since $\{m - n\upsilon_1 : n\}$ is uniformly bounded, Theorem 10 gives $n^{(d+2)/2}B(m - n\upsilon_1) \to 0$. Hence

$$n^{d/2}\left|\int_C f(x)\nu(dx + n\upsilon_1) - \frac{1}{\sqrt{(2\pi)^d\det[\upsilon_2]}}\int_C f(x)dx\right| \to 0$$

uniformly in $\theta \in \mathcal{K}$ and $C \in \mathfrak{C}$. □

Proof of Corollary 2.

In Theorem 7 take $f \in \mathcal{D}$ to be one on $|x| \leq b$, zero $|x| > b + 1$ and with bounded derivatives. □

7 Extended Perron–Frobenius

A matrix $M = \{m_{ij}\}_{p\times p}$, has nth power M^n with entries denoted by m_{ij}^n. The eigenvalues of M are the zeros of the characteristic polynomial $q(z) = \det[zI - M]$, which is of degree p. Denote the roots of $q(z)$ by $\rho_1, \ldots, \rho_p$ with the roots listed in order, so that

$$|\rho_1| \geq |\rho_2| \geq \ldots \geq |\rho_p|, \tag{7.16}$$

The eigenvalue ρ_1 of M is the maximum-modulus eigenvalue if $|\rho_1| > |\rho_2|$.

A non-negative square matrix $A = \{a_{ij}\}$ is called *positive regular* if all its entries are finite and, for some n, A^n has only strictly positive entries (see Lancaster and Tismenetsky (1985) or Seneta (1973). Let the entries $m_{ij}(\lambda)$ of the matrix $M(\lambda) = \{m_{ij}(\lambda)\}_{p\times p}$ be functions of $\lambda \in L$. Clearly the eigenvalues and their multiplicities all depend on λ.

Proof of Theorem 1.

Conclusions (i) and (ii) arise from routine applications of the implicit function theorem; details can be seen in Biggins and Rahimzadeh Sani (2004). Similar results have been obtained before, particularly for matrices of Laplace transforms. See, for example, Miller (1961, Theorem 1(a)), Lancaster and Tismenetsky (1985, Theorem 11.5.1), Ney and Numellin (1987, Theorem 4.1) and Kontoyiannis and Meyn (2003, Proposition 4.8(iii)). Now, by (i) there is an open set containing $\tilde{L}$ on which $|\rho_1| > |\rho_2|$, where $\rho = \rho_1$ — the eigenvalues here depend on λ but this is left implicit in the notation. For the rest of the proof λ will be confined to this open set.

The resolvent of $M(\lambda)$ is defined by

$$R(z) = \{r_{ij}(z)\}_{p\times p} = (I - zM(\lambda))^{-1},$$

which, for all i and j, has the expansion $r_{ij}(z) = \sum_{n=0}^{\infty} z^n m_{ij}^n(\lambda)$ when $z < ||M(\lambda)||^{-1}$, where $||\cdot||$ is a matrix norm, (see Lancaster and Tismenetsky (1985, Theorem 11.1.1)). For fixed $\tilde{\lambda} \in \tilde{L}$, let $B = B(\tilde{\lambda}, \tilde{\delta})$ and M_1 be such that for all $\lambda \in B$, $||M(\lambda)|| < M_1$. Now we take $|z| < 1/M_1$, $\lambda \in B$ and suppress λ in the notation.

Let $h(z) = \det[I - zM] = (1 - z\rho)\prod_{k=2}^{p}(1 - z\rho_k)$. Inverting $I - zM$ and rewriting using partial fractions

$$r_{ij}(z) = \frac{d_{ij}(z)}{(1 - z\rho)\prod_{k=2}^{p}(1 - z\rho_k)} = \frac{a_{ij}}{1 - z\rho} + \frac{\sum_{k=0}^{p-2} b_{k,ij} z^k}{\prod_{k=2}^{p}(1 - z\rho_k)}. \tag{7.17}$$

For fixed λ, $|b_{k,ij}/\rho^k|$ are bounded over i, j and k ($\leq p-2$), by C say. Then, expansion of (7.17) gives

$$r_{ij}(z) = \sum_{n=0}^{\infty}\left(a_{ij}\rho^n + \sum_{h=0}^{p-2} b_{h,ij} \sum_{k_2+\ldots+k_p=n-h} \rho_2^{k_2}\ldots\rho_p^{k_p}\right) z^n \tag{7.18}$$

and so

$$m_{ij}^n - a_{ij}\rho^n = \sum_{h=0}^{p-2} b_{h,ij} \sum_{k_2+\ldots+k_p=n-h} \rho_2^{k_2}\ldots\rho_p^{k_p}.$$

Hence, for $\alpha \in [|\rho_2|/|\rho|, 1)$

$$\left|\rho^{-n} m_{ij}^n - a_{ij}\right| \leq C(p-1)(n+1)^{p-1}\alpha^{n-p}. \tag{7.19}$$

Let A be the matrix with entries $\{a_{ij}\}$. Then $\rho^{-n}M^n \to A$ and so $MA = AM = \rho A$. Parts (i) and (ii) of the theorem now give $A = cvu$ for some scalar c; but $v = \rho^{-n}M^n v \to Av = cv$, using $uv = 1$, and so $c = 1$.

To translate (7.18) with $a_{ij} = v_i u_j$ into an asymptotic estimate of $M(\lambda)^n$, the boundedness of $b_{k,ij}$ as λ varies is now needed. Let $c_1(B)$ be the continuously differentiable functions on B. The function $h(z)$ is a polynomial in z with coefficients in $c_1(B)$ and $1/\rho$ is a simple root of this polynomial. Hence $h_1(z) = h(z)/(1 - \rho z)$ is a polynomial in z with coefficients in $c_1(B)$; the same is true of $d_{ij}(z)$. From (7.17) and $A = vu$

$$d_{ij}(z) = v_i u_j h_1(z) + \sum_{k=0}^{p-2} b_{k,ij} z^k (1 - z\rho).$$

Equating powers of z here shows that the $b_{k,ij}$ are in $c_1(B)$. Hence the supremum of $\left|b_{k,ij}(\lambda)/\rho(\lambda)^k\right|$ over i, j, k ($\leq p-2$) and λ in the closed ball of radius $\tilde{\delta}/2$ centred at $\tilde{\lambda}$ will be finite; denote this finite supremum by C.

Let 3ϵ be less than $(|\rho(\tilde{\lambda})| - |\rho_2(\tilde{\lambda})|)$ and small enough that balls of radius 3ϵ centred on the distinct eigenvalues of $M(\tilde{\lambda})$ are disjoint. Then any point within ϵ of one of $\rho_2(\tilde{\lambda}), \ldots, \rho_p(\tilde{\lambda})$ is smaller in magnitude than every point in the ϵ ball centred on $\rho(\tilde{\lambda})$. Let $\delta < \tilde{\delta}/2$ be small enough to ensure that the maximum distance between the eigenvalues at $\tilde{\lambda}$ and those at $\lambda \in B(\tilde{\lambda}, \delta)$, after a suitable pairing, is less than ϵ. Then, for $\lambda \in B(\tilde{\lambda}, \delta)$, and $j = 2, 3, \ldots, p$

$$\left|\frac{\rho_j(\lambda)}{\rho(\lambda)}\right| \leq \frac{|\rho_2(\tilde{\lambda})| + \epsilon}{\rho(\tilde{\lambda}) - \epsilon} \leq \frac{\rho(\tilde{\lambda}) - 2\epsilon}{\rho(\tilde{\lambda}) - \epsilon} = \alpha < 1.$$

With these definitions, the estimate (7.19) holds uniformly on $B(\tilde{\lambda}, \delta)$ and this implies (1.2). □

Although we do not need it here, it is worth noting the following result, which applies to suitable matrices of Fourier transforms. Its proof requires only obvious modifications of the discussion already given.

Theorem 11 *Theorem 1 remains true when 'analytic' is be replaced by '$q(\geq 1)$ times differentiable' and $\mathbb{C}^d$ is replaced by $\mathbb{R}^d$.*

Acknowledgements

Part of this work was carried out when ARS was on sabbatical at the University of Sheffield in 2003. ARS would like to thank the University of Sheffield for hosting him and the Mosaheb Institute of Mathematics of Teacher Training University for their financial support. We are grateful to Peter Hall for pointing us in the direction of Bhattacharya's work.

References

[1] Athreya, K.B. and Ney, P.E. (1972). *Branching Process.* Springer-Verlag, Berlin.

[2] Barral, J. (2001) Generalized vector multiplicative cascades. *Adv. Appl. Probab.* **33**, 874–895.

[3] Bhattacharya, R. N. (1972). Recent results on refinements of the central limit theorem. *Proc. Sixth Berkley Symp. Math. Statist. Prob.* **2**, 453–484. Univ. of California Press.

[4] Bhattacharya, R. N. (1977). Refinements of the multidimensional central limit theorem and applications. *Ann. Probab.* **5**, 1–27.

[5] Biggins, J. D. (1976). The first-and last-birth problems for a multitype age-dependent branching process. *Adv. Appl. Probab.* **8**, 446–459.

[6] Biggins, J. D. (1992). Uniform convergence of martingales in the branching random walk. *Ann. Probab.* **20**, 137–151.

[7] Biggins, J. D. (1996). How fast does a general branching random walk spread? In *Classical and Modern Branching processes*, (K.B. Athreya, P. Jagers, eds.). *IMA Volumes in Mathematics and its Applications* **84**, 19–40. Springer-Verlag, New York.

[8] Biggins, J.D. and Kyprianou, A.E. (2004). Measure change in multitype branching. *Adv. Appl. Probab.* **36,**, 544–581.

[9] Biggins, J. D. and Rahimzadeh Sani, A. (2004). Extended Perron–Frobenius results. Available at:
`http://www.shef.ac.uk/~st1jdb/epfr.html`
Preprint 539/04, Department of Probability and Statistics, University of Sheffield (February 2004).

[10] Bramson, M., Ney, P. and Tao, J. (1992). The population composition of a multitype branching random walk. *Ann. Appl. Probab.* **2,** 575–596.

[11] Chaganty, N. R. and Sethuraman, J. (1993). Strong large deviation and local limit theorems. *Ann. Probab.* **21**, 1671–1690.

[12] Hörmander, L. (1973). *An Introduction to Complex Analysis in Several Variables.* North-Holland, Amsterdam.

[13] Jensen J. L. *Saddlepoint Approximations*, Oxford University Press, Oxford, 1995.

[14] Kingman, J. F. C. (1961). A convexity property of positive matrices *Quart. Jnl. Math. Oxford.* **12**, 283–4.

[15] Kontoyiannis, Y. and Meyn S. P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.* **13**, 304–362.

[16] Kyprianou, A. and Rahimzadeh Sani A. (2001). Martingale convergence and the functional equation in the multi-type branching random walk. *Bernoulli.* **7**, 593–604.

[17] Lancaster, P. and Tismenetsky, M., *The Theory of Matrices, second edition with applications*, Academic Press, New York, 1985.

[18] Miller, H. D. (1961) A convexity property in the theory of random variables defined on a finite Markov chain, *Ann. Math. Statist.* **32**, 1260–1270.

[19] Mode, C.J. (1971). *Multitype Branching Processes.* New York, American Elsevier.

[20] Ney, P. and Numellin, E. (1987). Markov Additive Processes, I. *Ann. Probab.* **15**, 561–592.

[21] Seneta, E. (1973). *Non-negative Matrices.* Allen and Unwin, London.

[22] Stone, C. (1965). A local limit theorem for nonlattice multi-dimensional distribution functions. *Select. Trans. Math. Statist. Probab.* **36**, 546–551.

[23] Stone, C. (1967). On local and ratio limit theorems. *Proc. Fifth Berkley Symp. Math. Statist. Prob.* **2(ii)**, 217–224. Univ. of California Press.

[24] von Bahr, B. and Esseen, C.G. (1965). Inequalities for the r-th absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36**, 299–303.

[25] von Bahr, B. (1967). Multi-dimensional integral limit theorems. *Ark. Mat.* **7**, 71–88.